Guiding scissors to cut DNA

THE METHOD THAT BACTERIA USE TO KEEP VIRUSES IN THEIR PLACE WILL NOW HELP US TO TIDY THE HUMAN GENOME, WRITES **S ANANTHANARAYANAN**

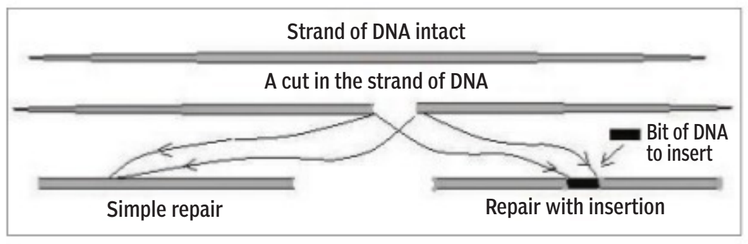
The human genome is a library of more than 20,000 genes that code for components of some two million proteins and are located over the length of the DNA, which has over three billion basic units. This genetic blueprint, which we carry in each of our 100 trillion cells, is a complex document and, in some parts, it directs the body to disease instead of health. We now know how to detect these problem genes, but doing anything about it is a different matter.

A solution has turned up based on something that was seen to be going on in the cells of bacteria. The solution has been refined into a technology to cut a DNA just where desired and then also to replace the portion that is cut with a modified segment when the DNA rejoins. The result is a flurry of research publications that report developments in DNA repair, over 100 papers in 2013, just after the method was discovered in the previous year, over 250 papers in 2014, and the numbers are increasing.

The discovery started with the observation that the chromosome, which is mostly DNA, of many bacteria contains a series of sequences that are repeated, interspersed with sequences that are derived from viruses. These repeats in the bacterial DNA, however, did not make much sense, till it was noticed that they usually occurred in the vicinity of another set of genes, of a kind that seemed to be involved in the process of DNA repair. It was then thought that the repeats were perhaps some kind of template acquired from encounters with viruses and later used for resisting the same viruses. The repeats were named *CRISPR*, a short form for Clusters of Regularly Interspaced Short Palindromic Repeats, and the associated genes were called *CAS* genes, for CRISPR Associated genes.

While the DNA are the construction formula of all the proteins that the organism would need, the whole DNA is never active at once. Instead, it is portions that represent the code for just a set of proteins that are extracted and utilised at each occasion of expression of genes and the creation of the different proteins that define the nature of a cell. The medium to transmit the partial code to the cells' engines that carry out the protein formation is the *RNA*, a variation of the DNA molecule structure.

Now, when a virus attacks a bacterium the identifying sequences of the virus get copied on to scraps of RNA that form during the encounter with the virus and these get stored in the CRISPR sequences. When the virus attacks again, it is not a passive action of taking and storing copies of parts of the virus DNA that happens, but the existing patterns of the virus DNA are copied quickly on to RNA, which then combine with proteins, or enzymes, that are formed from the CAS genes portion, enzymes of which an important group have been called *CAS9*.



CRISPR/CAS9 system

CAS9, equipped with the template for a portion of the virus, is then able to zero in on just the portion of the virus DNA that CRISPR has recorded and CAS9 does its wonderful thing of snipping the DNA just at that place. The virus, which consists mainly of its own DNA, is then destroyed and the bacterium stays healthy.

In 2012, scientist Jennifer Doudna of the University of California at Berkeley and her colleague, Emmanuelle Carpentier, and their team studied the process of this action and identified CAS9 as the enzyme that could cause a rupture in the two strands of the DNA molecule at the portion marked by the pattern of the RNA, called the *guide RNA*. When they had got down to how this happened in the bacteria, they wondered if the process could be mimicked with a given RNA to serve as the guide RNA to create a system that could sever the DNA at the place specified by the guide RNA that they had used. The team went on to do just this and then they devised a simple test to see if the target DNA had actually been cloven at the point they wanted.

They first generated a pair of guide RNA that matched specific parts of a known DNA molecule and then formed the cleaving enzymes by letting the RNA attach to CAS9. The target DNA was then incubated with the cleaving enzymes and then the DNA was examined with the help of a gel in which segments of DNA move with different speeds. The gel was able to separate two parts of the DNA and the sizes of the portions showed that the original DNA had been cut just where the cleavage was programmed by the guide RNA.

This discovery, Doudna says, was exciting as studies had shown for some time that cells had ways of repairing DNA damage. In such cases, when a DNA was separated, the two portions rejoined, either just as they had broken or with a small section added, or there could be the addition of a whole part of DNA that matched the two broken ends. Where the DNA rejoined in the ordinary way, there was often damage to the gene involved, which had its value in studying the properties or function of that gene. Where a portion was added, the genome now had additional genes, with their value in research. These features of DNA repair tempted scientists to hope for a way to break DNA at a place where there was an interest for repair, so that the introduction of a predetermined section could become possible.

The CRISPR and CAS9 system amounted to just that, and the DNA molecules could be divided accurately at the place where there was interest in carrying out a modification. After the nature of the DNA, as a chain built up of just four kinds of links, groups of atoms called A, G, T and C, in sequences that code for the production of all the proteins the body needs, was worked out by Crick and Watson in 1953, methods have been devised to navigate over the length of the chain and agents that are able to affect specific parts of the chain have been developed, etc. Specific genes that are implicated in specific diseases have been identified and there are methods to breed animals that lack that gene, or have a variation, to check whether this leads to remission of symptoms. The field is one of great promise, to treat genetic disease, generate crop varieties with greater yield or resistance to pests, bacteria that can change waste into useful substances, etc, but the processes to make modifications in the genome are painstaking and time consuming.

In this context, the CRISPR/CAS9 system, which provides the scientist and the industry the means to intervene directly at the DNA level, takes the possibilities of genetic engineering to another plane. It is now feasible to modify the genome of just any creature to any extent within a reasonably short period of time. This has its own dangers if used without care, but it could also enable bringing about very fast ecological changes to help the world cope with the challenges of warming in the coming decades.

THE WRITER CAN BE CONTACTED AT
response@simplescience.in



Jennifer Doudna and Emmanuelle Carpentier

EMBRYONIC DEVELOPMENT

TAPAN KUMAR MAITRA EXPLAINS THE ROLE PLAYED BY HOMEOTIC GENES

Some especially striking examples of coordinate gene regulation in eukaryotes involve an unusual class of genes known as homeotic genes. When mutations occur in one of these genes, a strange thing happens during embryonic development — one part of the body is replaced by a structure that normally occurs somewhere else. The discovery of homeotic genes can be traced back to the 1940s when Edward B Lewis discovered a cluster of *Drosophila* genes, the *bithorax gene complex,* in which certain mutations cause drastic developmental abnormalities such as the growth of an extra pair of wings.

Later, Thomas C Kaufman and his colleagues discovered a second group of genes that, when mutated, led to different but equally bizarre developmental changes — for example, causing legs to grow from the fly's head in place of antennae. This group of genes is the *antennapedia gene complex.* The bithorax and antennapedia genes are called homeotic genes because *homeo* means "alike" in Greek and mutations in these genes change one body segment of *Drosophila* to resemble another.

Although such phenomena might at first appear to be no more than oddities of nature, the growth of an extra pair of wings or the development of legs in the wrong place suggests that homeotic genes play key roles in controlling the formation of the body plan of developing embryos. Homeotic genes exert their influence over embryonic development by coding for a family of regulatory transcription factors that activate (or inhibit) the transcription of developmentally important genes by binding to specific DNA sequences.

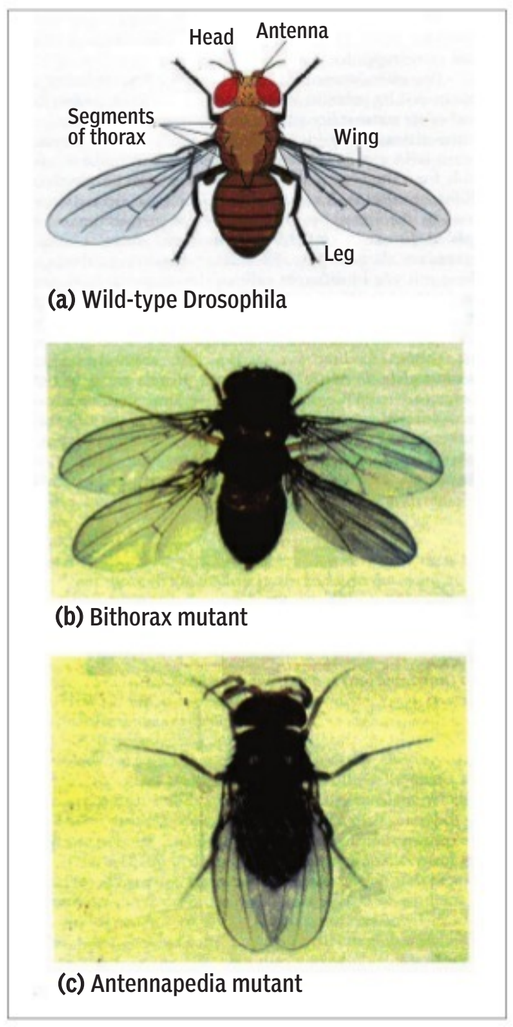
By binding to all copies of the appropriate DNA element in the genome, each homeotic transcription factor can influence the expression of dozens or even hundreds of genes in the growing embryo. The result of this coordinated gene regulation is the development of fundamental body characteristics such as appendage shape and location. Most homeotic genes control major developmental pathways, and in vertebrates they may also help regulate other important processes, such as histone production and antibody synthesis.

A clue to how homeotic genes work first emerged from the discovery that the bithorax and antennapedia genes each contain a similar, 180-bp sequence near their 3' end that resembles a comparable sequence found in other homeotic genes. Termed the *homeobox,* this DNA sequence codes for a stretch of 60 amino acids called a *homeodomain.* Homeobox sequences have been detected in more than 60 *Drosophila* genes and in a range of other organisms as diverse as sea squirts, frogs, mice and humans — an evolutionary distance of more than 500 million years. In fact, the C-terminal portion of the typical eukaryotic homeodomain even shows some homology with prokaryotic repressors.

The widespread occurrence of the homeobox suggests that the homeodomain for which it codes must perform an important function that has been highly conserved during evolution. Like the bacterial *lac* repressor, *trp* repressor and cAMP receptor proteins discussed earlier, the homeodomain contains a helix-turn-helix motif, suggesting that it functions in binding to DNA. This hypothesis has been confirmed by synthesising the homeodomain region of the transcription factor encoded by the antennapedia gene and showing that this 60-amino acid stretch binds to DNA in the identical sequence-specific manner as the intact transcription factor.

The role of the homeodomain helix-turn-helix motif in DNA binding has been further investigated in the homeotic gene called *bicoid,* which is required for establishing the proper anterior-posterior polarity during *Drosophila* development. The amino acid lysine is present in position number 9 of the recognition helix of the helix-turn-helix motif in the bicoid protein, whereas glutamine is present in the comparable position in the antennapedia protein. Normally the bicoid and antennapedia proteins function as transcription factors that bind to different DNA response elements.

However, if the lysine present at position 9 of the bicoid protein is mutated to glutamine, the altered bicoid protein binds to DNA sequences normally recognised by the antennapedia protein. Thus, not only does the helix-turn-helix motif of the homeodomain mediate DNA binding, but single amino acid differences within this motif can alter the DNA sequence to which the protein binds.



The wild-type Drosophila (a) has two wings and six legs extending from its three thoracic segments; mutations (b) in the bithorax gene complex convert the third thoracic segment to a second thoracic segment (the wing-producing segment), and an additional set of wings is formed; and a mutatioji (c) in the antennapedia gene complex causes legs to develop where the insect's antennae should be.

THE WRITER IS ASSOCIATE PROFESSOR, HEAD, DEPARTMENT OF BOTANY, ANANDA MOHAN COLLEGE, KOLKATA, AND ALSO FELLOW, BOTANICAL SOCIETY OF BENGAL, AND CAN BE CONTACTED AT
tapanmaitra59@yahoo.co.in

Immunity from interbreeding

KATE YANDELL REPORTS ON HOW NEANDERTHALS AND DENISOVANS CONTRIBUTED INNATE IMMUNE GENES TO MODERN HUMANS

According to two studies published on 7 January in *The American Journal of Human Genetics,* modern humans adopted innate immune genes responsible for recognising invading microbes from Neanderthals and Denisovans. Two teams, based in France and Germany, independently concluded that humans picked up some versions of a cluster of toll-like receptors by interbreeding with archaic hominin relatives.

"At least partially, Neanderthals may have harboured already adaptive mutations, mutations that rendered them more resistant to infections," said Luis Quintana-Murci, an evolutionary geneticist at the Pasteur Institute in Paris and a co-author of one of the new papers. Previous studies have shown that modern humans interbred with Neanderthals and Denisovans. For instance, Rasmus Nielsen, an evolutionary biologist at the University of California, Berkeley, and his colleagues showed that humans who migrated to Tibet likely picked up an allele controlling blood haemoglobin concentration from local Denisovans, allowing them to adapt to living at high altitudes. Another paper indicated that humans had picked up major histocompatibility genes from Denisovans and Neanderthals.

The authors of the two new studies approached the topic of ancient human evolution from different directions. Quintana-Murci and his colleagues decided to do a broad survey of innate immune genes and their variability among present-day humans around the world, using sequence data gathered through the 1,000 Genomes Project. The team demonstrated that innate immune genes have been under stronger-than-average selective pressures. Some innate genes are highly conserved, with little tolerance for variability. Other protein-coding genes have picked up adaptive mutations, mostly occurring within the last 6,000-13,000 years after humans transitioned from a hunter-gatherer to agricultural society. The resulting increase in density of human settlements, cohabitation with animals and increased exposure to sewage may have made humans easier targets for microbial disease, the researchers speculated.

Quintana-Murci and his colleagues also took advantage of a map of areas of the human genome where Neanderthal genes are present, showing that innate immune genes are generally more likely to have been borrowed from Neanderthals than genes coding other types of proteins. Specifically, they noted that 126 innate immune genes in present-day Europeans, Asians, or both groups were among the top five per cent of genes in the genome of each population most likely to have originated in Neanderthals.

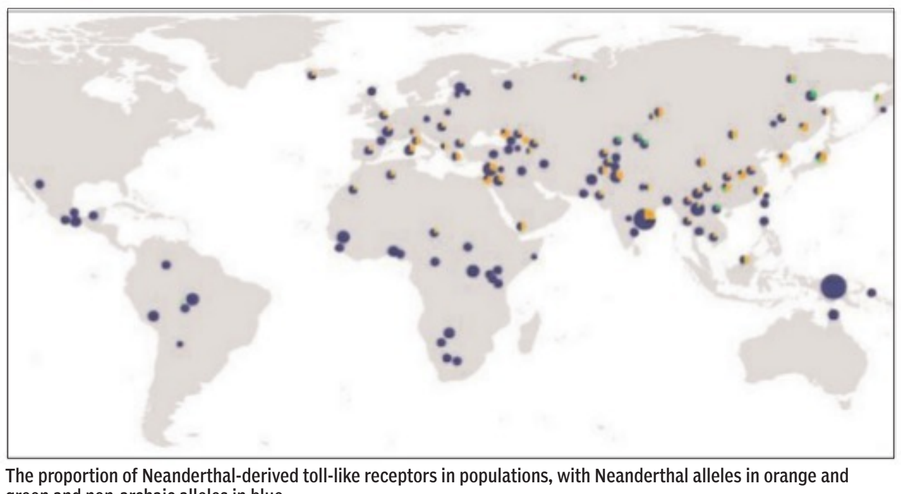
Another group led by Janet Kelso of the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, used both the same previously published Neanderthal introgression map that Quintana-Murci used and a second introgression map. They searched for borrowed regions of the genome that were especially long and common in present-day humans, eventually zeroing in TLR6, TLR10, and TLR1. These receptors, which detect conserved microbial proteins such as flagellin, are all encoded along the same segment of DNA on chromosome four.

By looking at 1,000 Genomes Project data, Kelso and her colleagues were able to identify seven distinct versions of the TLR cluster. They were able to match two of these versions to DNA from Neanderthals, and one version to DNA from Denisovans. She and her colleagues then attempted to figure out the functional differences between the Neanderthal and Denisovan versions of the TLR cluster and the versions that likely originated with modern humans who migrated from Africa to Europe and Asia later than these archaic hominids.

The changes in the Neanderthal and Denisovan TLR clusters do not lead to altered proteins. However, the researchers found that in white blood cells, the Neanderthal and Denisovan TLRs are more highly expressed than the non-borrowed human TLR clusters.

Kelso and her colleagues also did a survey of already-completed genome wide association studies, finding that present-day people who have the borrowed TLR clusters show lower levels of the bacterium *Helicobacter pylori* in their bloodstreams than people descended from humans that did not pick up TLR clusters from Neanderthals or Denisovans. People with the borrowed TLR clusters also tend to have elevated allergies to dust and pollen.

THE SCIENTIST


The proportion of Neanderthal-derived toll-like receptors in populations, with Neanderthal alleles in orange and green and non-archaic alleles in blue.

PLUS POINTS

Set for comeback

The warm glow of an old electric light bulb, which has been banned in many countries because it is so energy inefficient, may soon be making a comeback following a breakthrough in "light recycling". The problem with older light bulbs, which have not changed in their basic design since the days of Thomas Edison more than a century ago, is that they generate light by heating a thin filament of wire to very high temperatures.

This unfortunately means that about 95 per cent of the electrical energy used to power a light bulb is converted into heat and only five per cent into visible light, making it one of the most inefficient methods of lighting up a room.

However, a team of researchers from the Massachusetts Institute of Technology has devised a way of capturing the infrared heat released when the tungsten filament of an incandescent bulb is heated to about 2,700° Celsius, and converting this waste energy into visible light. They built a secondary structure around the incandescent filament which reflected the infrared light given off by the heated wire back to the filament where it is reabsorbed and then re-emitted as visible light. Because the secondary "filters" are not in direct physical contact with the hot filaments, they are not destroyed by the high temperatures.

The scientists have called the technology "light recycling" because it takes the useless wavelengths of infrared light, which no-body can see, and converts them to visible light. "It recycles the energy that would otherwise be wasted," said Marin Soljaci, professor of physics at MIT and a senior author of the study published in the journal *Nature Nanotechnology*.

STEVE CONNOR/THE INDEPENDENT

'Human-animal hybrids'

US scientists are growing human organs inside farm animals like sheep and pigs in research that could one day save the lives of people in need of organ donations. According to the *MIT Technology Review*, around 20 animals have been impregnated with human-animal hybrids at a number of universities in the last year. None of the hybrids have been born, but the technology used to create them is currently in development.

The "chimeras" are created by

injecting human stem cells into an animal embryo. These embryos are then implanted into a pig or a sheep. Researchers manipulate the stem cells inside the embryo in the hope that they will grow into a human organ of their choosing, which can be taken from the animal and implanted into a human.

People in need of organ transplants can wait years for the operation, due to demand far outstripping the number of donors, but thanks to the work of researchers at several institutions in America, human organs could one day be grown on demand inside animals, eliminating long waiting lists and saving lives.

DOUG BOLTON/THE INDEPENDENT

Barefoot engineers

Afghanistan has one of the lowest rates of electricity usage in the world and just 38 per cent of the population are connected to the

grid as poor infrastructure has been exacerbated by years of conflict. As part of a project, Norwegian Church Aid has introduced solar power to communities in four provinces previously cut off from electricity supplies: Bamyan, Daykundi, Faryab and Oruzgan.

The project is based on the barefoot movement in India aimed at empowering rural communities and developing skills by training local, often uneducated, people to adopt, manage and own sophisticated technologies without the need for external technicians. In 2005, 10 people travelled to India for training at the Barefoot College in installing solar systems. These master "barefoot solar engineers", in turn, have trained 84 engineers scattered across these provinces.

SCIDEV.NET